DOCUMENT RESUME

ED 413 763 FL 024 835

AUTHOR	Taura, Hideyuki
TITLE	Bilingual Measurement.
PUB DATE	1996-09-00
NOTE	13p.
PUB TYPE	Guides - Non-Classroom (055) -- Reports - Evaluative (142)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Bilingualism; Comparative Analysis; Ethnography; Foreign Countries; Language Dominance; Language Proficiency; Language Research; *Language Role; Language Skills; *Language Tests; *Measurement Techniques; Psychometrics; Qualitative Research; Research Methodology; *Testing

ABSTRACT

An examination of measures currently used to assess degree of bilingualism in individuals looks at the problems inherent in them, ideal measures, and how to use available measures. It is suggested that definition of bilinguality is a central cause for measurement problems, since a bilingual is assumed to be the sum of two monolinguals, without regard to the specific competence of a bilingual; reconceptualization of bilinguality is recommended. Analysis then looks at two types of measurement, formal and informal. Formal measurement is divided into traditional (focusing on grammar, pronunciation, and vocabulary) and behavioral (language balance and dominance) measures. Problems with traditional measures include lack of measures for native speakers, loss of test validity through translation into another language, examiner unfamiliarity with the language or culture, and minimal measurement of communicative competence. Behavioral measures, which assume that language tasks would elicit similar performance in both languages, have questionable validity. Informal measurements include language background questionnaires and ethnographic data collection. Both are found to have methodological limitations, principally lack of precision and possible researcher bias. It is recommended that researchers use a combination of currently available measures, depending on the kind of information needed. Contains 17 references. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 413 763

Bilingual measurement

Hideyuki Taura

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
Hideyuki Taura
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

FL024835

1 INTRODUCTION

The purpose of this essay is to show (1) the problems inherent in the bilingual measurements currently used, (2) ideally desirable measurements, and (3) how to best use the available bilingual measurements from a practical point of view.

The definition of 'bilinguality' and 'bilingualism' is considered first, since these terms seem to be used interchangeably. According to Hamers and Blanc (1989, p6) 'bilinguality is the psychological state of an individual who has access to more than one linguistic code as a means of social communication while the concept of bilingualism includes that of bilinguality but refers equally to the state of a linguistic community in which two languages are in contact'. Following this definition, this essay uses the term 'bilinguality' to refer to a psychological state and 'bilingualism' as a state of a linguistic community.

2 PROBLEMS INHERENT IN BILINGUAL MEASUREMENTS

Firstly, the existing bilingual measurements are overviewed to explore their significance and limitations.

2.1 Definition of bilinguality

The main cause for current bilingual measurement problems seem to be the definition of bilinguality. One extreme definition is put forward by Bloomfield (1933, cited in Romaine 1995, p11) as 'native-like control of two languages'. At the other extreme end of the continuum is Diebold (1964, cited in Romaine 1995, p11) whose definition is minimal, when he refers to 'incipient bilinguals'. This indeterminate definition of bilinguality leads researchers to a further complicated issue of determining the degree of bilinguality when they attempt to measure language proficiencies of bilinguals, regardless of researchers' position on the definition continuum. The most frequently used technique for measuring bilinguality, for instance, would be to

take measures in each of the bilingual's two languages and to compare them with monolingual standards. This methodology takes the stance that a bilingual is the sum of two monolinguals, and does not measure features specific to bilinguals such as code-switching. Drawing attention to this inadequacy of monolingual measures on bilinguals, Lavandera (1978, cited in Hamers & Blanc, 1989, p21; Romaine, 1995, p19) argue that it is an urgent task for researchers to design a measurement test that captures the bilingual's specific competence. This is based on the assumption that bilinguals' competence can only be measured when both languages are fully exploited and bilingual-specific strategies and competence are studied. This issue also raises a validity question of comparing the bilingual competence in each language with a monolingual's (Grosjean, 1985, 1989; Baker, 1993; Romaine, 1995). Thus before operationalizing or measuring the bilinguality, the concept of bilinguality must be conceptualized or defined so that all researchers can work from the same basic framework.

2.2. Traditional measurements

There have been two types of measurements - formal and informal. The formal measurements have subcategories of traditional and behavioural measurements. First of all, the traditional measurements are examined.

Traditional measurements have tested bilingual L1 and L2 linguistic proficiencies in smaller components such as grammar, pronunciation, vocabulary etc., and compared the results with monolingual norms. This involves at least three problems. Firstly, even competence in native speakers varies greatly, therefore 'procedures to measure native competence have not been developed' (Valdes & Figueroa, 1994, p67). The issue is what comparisons can be made ? Secondly, there is doubt cast on the total of the results in each component test as representing a bilingual's bilinguality, since 'competence is not divisible into isolated components' (Romaine, 1995, p18). The third problem is related to the norms with which the results

are compared. There exist some normative data in each developmental (age stage) phase, at least in English, from the monolingual first language acquisition studies for comparison. However as for other minor languages there is hardly any developmental data in L1 acquisition, let alone in L2 studies (Stokes & Duncan, 1989, p115). Therefore data accumulation on developmental features in each language is urgently needed to enable traditional methods to be used as a comparison. Furthermore, from the standpoint that bilinguals are not the sum of two monolinguals, normative data, in the true sense, taken from bilinguals must be accumulated in their individual languages, as Long (1990) contends that lexical items familiar with monolinguals may not be so within bilingual-bicultural family situations.

Secondly, it is pointed out that traditional tests may lose validity when they are translated directly into other languages and tested upon a different population (Stokes and Duncan, 1989). In order to obtain an overall picture of a bilingual's linguistic proficiency in both languages, tests directly translated from the English version have often been used, due to the lack of proficiency tests in certain languages. Stokes and Duncan (1989) argue that a standardized test targeted, for instance, at North American white, middle-class children would not be valid when they are tested on working-class children from the urban areas of England, let alone on children where language and cultures are totally different. Similar comment is made by Long (1990, p308) 'many standardized language tests are significantly biased against bilingual-bicultural children'. Therefore standardized (or its translation version) tests must always involve linguistic and cultural modification and restandardization, depending on the target population, to eliminate cultural bias.

Thirdly, there are problems with reference to examiners. As Long (1994) claims, examiners might misinterpret the data because they are not familiar with the particular language or culture under question. They might jump to a conclusion based on only a limited amount of

data.

Lastly, traditional tests may be good tools in terms of taking a relatively short time to administer as well as facilitating 'feedback to the teacher that directly leads to action' (Baker, 1993, p21). Each test, however, sheds light on only one aspect of bilingual competence and does not indicate how each skill is integrated. Tests such as cloze-tests are invented to capture the unitary aspects of bilingual proficiencies in that examinees must utilize grammatical, syntactical, lexical, semantic, collocational, contextual, situational, and logical clues (Oller, 1979; Heaton, 1988; Piper, 1983). Even cloze-tests, however, cannot measure some aspects of communicative competence such as knowledge of illocutionary rules or appropriate use of linguistic rules in communication settings. It is this communicative competence that traditional tests have particularly failed to measure. Since 'reducing everyday language competence to tests of specific skill is like measuring Michelangelo's art solely by its range of colours' (Baker, 1993, p26), researchers such as Abundarham (1980, cited in Stokes & Duncan, 1989, p118), Skehan (1988) and Grosjean (1985, 1989) propose that the best bilingual measurement method would be to see how bilinguals perform in both languages in a range of real communicative situations. Canale and Swain (1980) and Bachman (1990, cited in Baker, 1993, p31) suggest, holding that language competence is composed of 2 subcomponents - organizational (grammatical and textual) competence and pragmatic(illocutionary and sociolinguistic) competence, that bilinguals should be measured in the integration of all these skills. Towards this ideal, the US Foreign Service oral interview has been invented and used extensively with some resultant insight in the communicative competence measurements. Nonetheless, this is not without any criticism as to its validity - whether reality is reflected in such an unnatural, artificial situation however trained the interviewers might be. Furthermore Kessler (1984, cited in Hoffmann, 1991, p150) argues 'the acquisition of communicative competence in two languages must take into account the interaction between two language system',

and mentions four major interrelated factors; (1) context-bound home language vs. decontextualized language at school, (2) relationship between bilinguals and their interlocutors, (3) affective variables such as bilingual's emotional bonds with speakers of his/her two languages, and attitudes towards two languages, and (4) the need to use L1 and L2. Accordingly, bilinguality of a bilingual is constantly changing, depending on the environment in which he/she is, and the results obtained in the formal tests could only represent the bilingual's tentative and partial proficiency in both languages. It would be difficult to measure communicative competence in an unbiased, comprehensive, valid, and reliable way, however there is the need to devise a method to measure it.

2.3. Behavioural measurements

The difficulty of measuring bilingual competence by means of traditional language tests have led researchers to design psychometric tests directly comparing the two languages and assessing their language balance and dominance. These tests are devised on the assumption that a balanced bilingual's performance would be the same whatever the language used in performing the task. Outlines of some of these tests are discussed in terms of their validity and reliability below.

The first psychometric measurement is a reaction-time measurement test, where a bilingual is assessed as 'balanced' if the reaction time for a L1 word presented through a tachistoscope is equal to that for its translation equivalent in L2. Although the results on this method are found to correlate highly with those of traditional linguistic measurements (Lambert, Havelka, and Gardner, 1959, cited in Hamers & Blanc, 1989, p17), it is problematic because of the frequency distribution of the translation equivalents.

The second category is a word-completion task, in which bilinguals are presented with the beginning of potential words and

asked to produce as many words as possible in both languages. The results gained on this test is also known to have high correlation with the traditional measures of language competence (Hamers & Blanc, 1989, p18), the problem is obvious - this method is impossible unless the two languages share a common script.

The third type of test is carried out through reading aloud a list of cross-language ambiguous words. Despite of its high correlation with other types of bilingual balance measures, this method is problematic firstly because of its limited use of languages which share a similar lexicon and secondly the different frequency distributions in both languages.

Thus behavioural measures may have the advantage of being simple, easy to conduct and comparing two languages directly. However the principal assumption on which this type of measure is based, that is, some tasks would elicit the same performance in both languages from balanced bilinguals, seems elusive. Together with evidence such as Fishman's (1971, cited in Romaine, 1995, p19) that 'bilinguals are rarely equally fluent in both languages about all possible topics', it would be, therefore, desirable to combine behavioural measures with other measures.

2.4. Informal measurements

Formal measurement tests are conducted in clinical or unnaturalistic environments. This has raised some questions such as whether these data represent reality because they are collected where bilinguals are under some psychological pressure, or whether some tests may be examining their cognitive skills (Hoffmann, 1991), or at worst, test-taking skills (Cummins & Swain, 1986, p141) instead of bilinguality. In response to this issue, two different kinds of informal data-gathering have been constructed.

The first type of data-gathering is a language background questionnaire, which asks bilinguals to provide information on various aspects of their bilinguality such as the age and context of acquisition of the both languages. This information is useful in that it provides clues to the type of bilinguality or the use of each language, the degree of literacy, etc. Self-ratings or teachers judgements on language proficiencies in both languages on a seven-point scale (Macnamara, 1967) are used as well, in order to obtain evaluation on bilingual competence. Despite their usefulness, it should be kept in mind that the data through these three methods are all filtered by many variables such as raters' subjectivity (Baker, 1993), attitudes towards languages and status of the languages (Romaine, 1995). Therefore it is desirable to use these methods in combination with other measures to enhance reliability.

The second type of informal measurement is ethnographical data collecting. Unable to capture the competence specific to bilinguals such as code-switching, and incapable of assessing bilingual children's competence in minority groups, researchers have turned from the framework of traditional or behavioural psychometric measurements to ethnographic and sociolinguistic approaches. In doing so, researchers have taken two approaches - the macro-sociological and micro-sociological approaches (Hamers & Blanc, 1989, p25). The macro-sociological approach is operated in a form of censuses or surveys with large samples. While this approach provides statistical information on linguistic assimilation and diversity, the data do not show how intergroup and interpersonal communications are taking place in a multicultural setting. Another methodological limitation is that it does not allow researchers to use 'sensitive instruments and refined quantification, such as precise and detailed questions or the recording of actual language behaviour' (Hamers & Blanc, 1989, p28). The micro-sociological, or the ethnographical approach, on the other hand, enables researchers to use in-depth methods of data collection by

anthropological observation through interviews, diary-keeping, audio/video-recording, and observation, and provide rich and detailed analysis, although it is limited to small numbers and to description of phenomenon. Thus either approach at the societal level is still on its way to sophistication in methodology to measure what researchers are attempting to study.

3 IDEAL BILINGUAL MEASUREMENTS

After seeing the problems inherent in the currently used measurements, one would undoubtedly notice that the concept of bilinguality competence must be operationalized first. Otherwise conclusions which we draw from various tests are 'only as good as our measurements - operationalization of our theoretical constructs' (Cummins & Swain, 1986, p114). Since the bilingual is more than the sum of two monolinguals, bilingual specificity must be included in the measurements as well. The bilingual measurements also have to explore both linguistic and communicative competence. As for data-collection, data should be gathered formally (elicitation method) to compare with other bilinguals in the all areas of phonology, syntax, semantics, and non-verbal skills. At the same time, naturalistic data should be collected informally to grasp a rich and detailed picture of bilinguality. Assessment of bilinguality should take into consideration such variables as affective variables, bilingual types (sequential, simultaneous, etc.), the age of exposure to languages, socio-economic and educational background, and levels of intelligence. In order to assess the collected data in comparison with other bilinguals not with monolinguals, normative data should be accumulated. The normative data should cover a whole range of different target population in that Japanese-English bilinguals, for instance, should be assessed by the norms taken from the same population, not from French-English bilinguals, nor monolingual Japanese. The normative data should also provide developmental changes in individual languages in bilinguals, just as

English language has its developmental data acquired by monolingual children as L1. Finally examiners must be familiar with the two languages and cultures to validly assess the resultant data.

4 PRACTICAL MEASUREMENTS FOR INTERIM

Conducting bilingual measurements as described above is ideal, but it would be very difficult, if not impossible, in terms of practicality, particularly time-wise. Therefore the best method which researchers could employ for the time being would be to use a combination of existing measures, depending on exactly what aspect of bilinguality they are examining. The minimum requirement would be a test in both languages (no translation equivalent) which investigates both linguistic and communicative proficiencies. Due to the lack of bilingual norms up to data, the researcher can only compare the data to monolingual norms. In this comparison, the researcher should accumulate his/her own data towards the establishment of bilingual norms. As for bilingual specificity, which is an important part of bilinguality and should be tested, it is up to individual researcher whether it is included or not, since (1) there is not good tools to measure it, (2) the existing methods take time, and (3) above all the notion itself is not theoretically operationalized.

REFERENCES

Baker, C. (1993). *Foundations of Bilingual Education and Bilingualism.* U.K: Multilingual Matters.

Canale, M. and **Swain, M.** (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics* **1**, 1-47.

Cummins, J. and **Swain, M.** (1986). *Bilingualism in Education* . London:Longman.

Grosjean, F. (1985). The bilingual as a competent but specific speaker-hearer. *Journal of Multilingual and Multicultural Development.* **6** (6), 467-477.

Grosjean, F. (1989). Neurolinguists, Beware! The Bilingual is Not Two Monolinguals in One Person. *Brain and Language*, **36**, 3-15.

Hamers, J.F. and **Blanc, M. H.** (1989). *Bilinguality and Bilingualism.* Cambridge: Cambridge University Press.

Heaton, J.B. (1988). *Writing English Language Tests.* London: Longman.

Hoffmann, C. (1991). *In Introduction to Bilingualism.* Harlow: Longman.

Long, S. (1990). Language and bilingual-bicultural children. In V.A. Reed (ed.) *An Introduction to children with language disorders* (2nd ed). New York: Macmillan College Publishing Co.

Macnamara, J. (1967) The bilingual's linguistic performance: A psychological overview. *Journal of Social Issues* **23**, 58-77.

Oller, J. W. (1979) *Language Test at School.* London: Longman.

Paivio, A. (1981). *Psychology of language*. N.J.: Prentice-Hall.

Piper, A. (1983). A comparison of the Cloze and C-test as placement test items. *British Journal of Language Teaching*, **11**, 45-50.

Romaine, S. (1995). *Bilingualism*. 2nd edition. Oxford: Blackwell Publishers Ltd.

Skehan, P. (1988). *Language Testing, Language Teaching*, **21**, 1-13.

Stokes, J. and **Duncan, D.** (1989). Linguistic assessment procedures for bilingual children. In D. Duncan (ed.) *Working with bilingual language disability*. London: Chapman & Hall.

Valdes, G. and **Figueroa, R.A.** (1994). *Bilingualism and Testing: A Special Case of Bias*. New Jersey: Ablex Publishing Corporation.